FORM PTO-1390 (REV. 5-93)

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE

JC17 Rec'd PCT/PTO 16 APR 2001

ATTORNEY'S DOCKET NUMBER
11150/30

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)
09/807638

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED:
PCT/EP99/06476	**3 September 1999 (03.09.1999)**	**16 October 1998 (16.10.1998)** **25 February 1999 (25.02.1999)**

TITLE OF INVENTION
METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR MESSAGES, USING SPEECH

APPLICANT(S) FOR DO/EO/US
Ingo BOECKMANN; Holger EBERT; Matthias HEIMERMANN

Applicant(s) herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) immediately rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☒ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)) (unsigned)
10. ☒ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A **FIRST** preliminary amendment.
    ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
14. ☒ A substitute specification and marked-up version.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information: International Search Report (translated), Form PCT/RO/101, International Preliminary Examination Report

Express Mail No.:EL594612921US

U.S. APPLICATION NO (if known, see 37 C.F.R. 1.5)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
09/807638	PCT/EP99/06476	11150/30

17. ☒ The following fees are submitted:

	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5)):		
Search Report has been prepared by the EPO or JPO $860.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) $690.00		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) $710.00		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO $1,000.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) $100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 860	
Surcharge of $130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$	

Claims	Number Filed	Number Extra	Rate		
Total Claims	11 - 20 =	0	X $18.00	$0	
Independent Claims	2 - 3 =	0	X $80.00	$0	
Multiple dependent claim(s) (if applicable)			+ $270.00	$	
TOTAL OF ABOVE CALCULATIONS =				$860	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).				$	
SUBTOTAL =				$860	
Processing fee of $130.00 for furnishing the English translation later the ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)). +				$	
TOTAL NATIONAL FEE =				$860	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property +				$	
TOTAL FEES ENCLOSED =				$860	
				Amount to be refunded: $	
				charged: $	

a. ☐ A check in the amount of $_____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 11-0600 in the amount of $860.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 11-0600. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Kenyon & Kenyon

26646
PATENT TRADEMARK OFFICE

SIGNATURE

Richard L. Mayer, Reg. No. 22,490
NAME

4/16/01
DATE

[11150/30]

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s)	:	Ingo BOECKMANN et al.
Serial No.	:	To Be Assigned
Filed	:	Herewith
For	:	METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR MESSAGES, USING SPEECH
Examiner	:	To Be Assigned
Art Unit	:	To Be Assigned

Assistant Commissioner for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

S I R:

Kindly amend the above-captioned application before examination, as set forth below.

IN THE SPECIFICATION AND ABSTRACT:

In accordance with 37 C.F.R. § 1.121(b)(3), a Substitute Specification (including the Abstract, but without claims) accompanies this response. It is respectfully requested that the Substitute Specification (including Abstract) be entered to replace the Specification of record.

IN THE CLAIMS:

On the first page of the claims, first line, change "Claims" to --WHAT IS CLAIMED IS:--.

Please cancel, without prejudice, claims 1 to 10 in the underlying PCT application. Please also cancel without prejudice, claims 1 to 10 in the annex to the International Preliminary Examination Report.

Please add the following new claims:

--11. (New) A method for outputting at least one of information and status messages of at least one electrical device using speech, comprising the steps of:

storing the at least one of information and status messages relating to a voice output in a speech memory;

selectively reading the at least one of information and status messages by a processing device; and

outputting the at least one of information and status messages on an output device using an intonation in accordance with a relevance.

12. (New) The method according to claim 11, wherein the output device includes a loudspeaker.

13. (New) The method according to claim 11, wherein the at least one of information and status messages requiring immediate action are output in the outputting step using a command intonation.

14. (New) The method according to claim 11, wherein the at least one of information and status messages requiring immediate action are output in the outputting step at a high volume.

15. (New) The method according to claim 11, wherein the at least one of information and status messages requiring immediate action are output in the outputting step in a harsh manner.

16. (New) The method according to claim 11, wherein the at least one of information and status messages are stored in the speech memory in accordance with a plurality of speaking voices, the method further comprising the step of changing the speaking voice for the at least one of information and status messages requiring immediate action.

17. (New) The method according to claim 11, further comprising the step of increasing the intonation and a connotation of the at least one of information and status messages requiring immediate action in accordance with importance.

18. (New) The method according to claim 11, further comprising the step of varying the intonation with a decreasing connotation for the at least one of information and status messages not requiring immediate action.

19. (New) The method according to claim 11, further comprising the step of controlling the at least one electrical device using voice recognition.

20. (New) The method according to claim 11, wherein the storing step includes the substep of storing in the speech memory a plurality of alternatives of the at least one of information and status messages, and wherein the outputting step includes the substeps of:

successively outputting the alternatives of the at least one of information and status messages in response to a failure to interact with the at least one of information and status messages until an interaction occurs; and

changing a dialog-communication level in response to a failure to interact with a last of the successive alternatives of the at least one of information and status messages.

21. (New) A device for outputting at least one of information and status messages of at least one electrical device using speech, comprising:

a speech memory configured to store data relating to a voice output of the at least one of information and status messages;

a processing device; and

an acoustic output device, the at least one of information and status messages being selectively output on the output device using an intonation in accordance with a relevance.--.

REMARKS

This Preliminary Amendment cancels, without prejudice, claims 1 to 10 in the underlying PCT Application No. PCT/EP99/06476. This Preliminary Amendment further cancels, without prejudice, claims 1 to 10 in the annex to the International Preliminary Examination Report and adds new claims 11 to 21. The new claims,

inter alia, conform the claims to U.S. Patent and Trademark Office rules and do not add any new matter to the application.

The above amendments to the specification and the abstract conform the same to U.S. Patent and Trademark Office rules and do not introduce any new matter into the application.

The underlying PCT Application No. PCT/EP99/06476 includes an International Search Report, dated March 23, 2000, a copy of which is included. The Search Report includes a list of documents that were considered by the Examiner in the underlying PCT application.

The underlying PCT Application No. PCT/EP99/06476 also includes an International Preliminary Examination Report dated September 20, 2000. An English translation of the International Preliminary Examination Report and annex thereto is included herewith.

It is respectfully submitted that the subject matter of the present application is new, non-obvious and useful. Prompt consideration and allowance of the application are respectfully requested.

Respectfully submitted,

KENYON & KENYON

Dated: 4/16/01

By: [signature]

Richard L. Mayer
Reg. No. 22,490

One Broadway
New York, New York 10004
(212) 425-7200

[11150/30]

METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR STATUS MESSAGES, USING SPEECH

FIELD OF THE INVENTION

The present invention relates to a method and a device for outputting information and/or status messages of at least one electrical device, using speech.

BACKGROUND INFORMATION

Methods and devices of this type are generally used in so-called interactive voice-communication systems or voice-controlled systems for, e.g. vehicles, computers, robots, machines, equipment, etc.

In general, an interactive voice-communication system (SDS) can essentially be reduced to the following components:

- Voice recognition system, which compares an orally input command ("voice command") to other allowed voice commands, and decides which command, in all probability, was orally input;
- Voice output, which outputs the voice commands and signal tones necessary for prompting the user, and possibly acknowledges the recognition result;
- Dialog and sequencing control, in order to explain to the user which type of input is expected, to check if the input is consistent with the prompt and the current status of the application, and to trigger the resulting action in the application (for example, the device to be controlled);
- Control interface as an interface to the application: Hidden behind it are hardware and software modules for controlling various actuators and computers, which contain the application; and
- Application that is controlled by speech: For example, it can be an ordering or information system, a CAE workstation, or a wheel chair for the disabled.

For example, such a voice recognition system is described in German Published Patent Application No. 195 33 541. To increase the acceptance of such man-machine dialog, synonymous words or various pronunciations for the commands are used, or the words are rearranged in the commands. For example, "larger radius when turning left" can alternatively be expressed here as "when turning left, larger radius". In addition, a multilingual, interactive communication system independent of the speaker can be set up by expanding the memory, it being possible to alternatively switch between the interactive communication systems of various languages. In addition, ellipses may be used, i.e., dispensing with the repetition of complete command sentences, and instead using commands such as "higher", "sharper", or "further", the voice recognition system then assigning these to the preceding commands. In response to uncertain recognition, the voice recognition system can also pose questions such as "Excuse me?", "Please repeat that", or "What else?", or issue specific suggestions such as "Louder, please". All of these measures are used to avoid monotonic communication and to have the dialog more closely approximate human-to-human communication. To improve the communication, the voice system is coupled to an optical display medium, on which the recognized commands are indicated for control purposes. Furthermore, the optical display medium allows the display of functions from the target device which are set in response to the voice command; and/or the display of various functions/alternatives, which can subsequently be set or selected by a voice command. A disadvantage of this device and the method implemented thereby is that, despite the given improvements, the voice output tires the user due to its monotony, so that his or her reaction time is too slow during events requiring immediate action. An additional problem is that, in response to recognition difficulties, the voice recognition systems perform an endless loop and issue the user the same prompt again and again, so that the workflow is interrupted.

Therefore, it is an object of the present invention is based on the engineering to provide a method and a device for outputting information and/or status messages, using speech, in which the attentiveness of the user is improved.

SUMMARY

The above and other beneficial objects of the present invention are achieved by providing a device and method as described herein

By using different intonations, the attention of the user is immediately obtained while the speech is being output, so that the reaction time for performing the requested instruction is considerably reduced. In the case of instructions requiring immediate action, the status messages have a command intonation.

To further increase the attention span, and the differentiation of instructions requiring immediate action, the volume of the voice output may be increased for instructions requiring immediate action, and/or these instructions may be inserted in a particularly harsh or abrupt manner.

In addition, the voice recognition system may be designed to be multilingual, so that, for example, one may choose between a man's voice and a woman's voice. One of these voices are selected by the system, for instructions requiring immediate action, and the other is selected by the system, for information or status messages not requiring immediate action.

To ensure the workflow, the voice recognition system is only activated by actuating a "Push to talk" (PTT) switch, the dialog-communication level being changed in the absence of a valid interaction. To increase the recognition reliability and improve the user prompting, individual commands may be saved in various, alternative output forms, which are then

successively output in response to an invalid interaction. The dialog-communication level is only changed when a valid interaction does not ensue in response to all of the command forms. To avoid monotony, the sequence of the output may be permutated by a random-number generator.

The basis of the present invention is to use the manner in which speech is output to the motor vehicle driver, in order to create an emotion that causes one to act in accordance with the situation.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flowchart illustrating a method for automatically controlling at least on device using voice recognition according to the present invention.

Figure 2a illustrates a graph of a potential danger during an interaction that does not require immediate action.

Figure 2b illustrates a denotation graph corresponding to Figure 2a

Figure 2c illustrates an intonation graph corresponding to Figure 2a

Figure 2d illustrates a connotation graph corresponding to Figure 2a

Figure 3a illustrates a graph of a potential danger during an interaction that requires immediate action.

Figure 3b illustrates a denotation graph corresponding to Figure 3a.

Figure 3c illustrates an intonation graph corresponding to Figure 3a.

Figure 3d illustrates a connotation graph corresponding to Figure 3a.

DETAILED DESCRIPTION

The voice recognition system is activated by actuating a PTT switch. For clarity, the voice output of the voice recognition system is subdivided into commands KOM and prompts Auff which, in reality, may be identical. Hereinafter, commands KOM are to be understood as a direct instruction to act, such as "BRAKE" or "TURN ON LIGHT", whereas prompts Auff request an interaction in the form of an input, such as "Please specify desired temperature in degrees C."

If the voice recognition system now generates a command KOM, then this command KOM is subdivided according to whether it is an instruction requiring immediate action or an instruction not requiring immediate action. More simply, instructions requiring immediate action are commands KOM, which call for the action to be performed quickly. An example of this is command KOM "Brake", when an ADR system or a precrash sensory system has detected a collision object. Examples of instructions not requiring immediate action include commands KOM of a navigation system. In this context, instructions requiring immediate action are inserted in time t_1, with command-intonation voice S1 and high volume L1, in a harsh and abrupt manner, in order to produce a high degree of attentiveness in the user. However, instructions not requiring immediate action are inserted softly, at low volume L2 and normal intonation S2.

As a rule, time is not a critical factor in the case of prompts Auff, so that, in this case, good user prompting is of concern. For this purpose, n different alternatives of a prompt Auff may be stored in the speech memory. For example, the alternatives may be different emphases, pronunciations, word rearrangements, or synonymous terms. After acoustically outputting the first alternative, the voice recognition system

waits for a predetermined period of time for an interaction. If no interaction or an invalid interaction occurs within this time period, then the voice recognition system repeats the prompt, using the subsequent alternative up to the nth alternative, if necessary. If a valid interaction occurs, then this request is performed and, if necessary, a new prompt Auff is output. But if no valid interaction occurs in response to the nth alternative, then the system switches to another dialog-communication level DKE, in order to ensure the workflow. For example, new dialog-communication level DKE may then be a selection list, which is displayed on the trip-computer monitor, and from which the user may select a corresponding menu.

Figures 2a-d schematically represent the conditions for an instruction not requiring immediate action, such as an information prompt for a navigation system. In Fig. 2a, the importance of the interaction is plotted over time. Instructions for action are output at times t_0 - t_2, and it is assumed that there was no reaction to each preceding prompt. Since a missing input in the navigation system only results in the inoperability of comfort components, which are also not necessarily desired by the motor vehicle driver, the importance does not change over time. The information regarding the content of the command, or the so-called denotation, i.e. the input request, also remains constant over time; as illustrated in Fig. 2b. At time t_0, the motor vehicle driver may be prompted, "Please input your desired destination now." This prompt is issued, using a certain intonation I_1 and a certain connotation K1, which are illustrated in Figs. 2c and 2d. If nothing is input, then the system does not know the reason for omission, e.g., if the motor vehicle driver did not hear the request or deliberately intended not to perform it. Therefore, the prompt, "Would you like to input a destination," is issued again at time t_1, using a stronger intonation I_2, in order to improve the possibility of it being perceived. However, connotation level K2 decreases. If, in

response, nothing is input again, then the system may certainly determine that the motor vehicle driver does not wish to do this. To avoid annoying the motor vehicle driver with constant repetition, a prompt such as "If you do not wish to input a destination, I will now turn myself off" is then issued one last time, at time t_2. This last prompt is output, using a very low intonation I_3, and it just has a low connotation. As illustrated in Fig. 2d, the connotation forms an anticlimax, i.e., a transition from a strong to a weak expression, whereas a certain variation occurs in the intonation, in order to counteract monotony.

In contrast, Figs. 3a to 3d illustrate represent a situation in which the importance of the interaction increases over time, until action is finally required. For example, the motor vehicle travels on a motorway at a speed greater than an allowed speed, while maintaining the safety distance behind a motor vehicle. At time t_0, the system issues an action instruction to the motor vehicle driver, e.g., in the form of "Please adjust your speed." The action instruction has a low intonation degree I_1 and a correspondingly low connotation level K1 since the motor vehicle driver is indeed acting illegally, but no immediate danger exists. In addition, it is now assumed that the motor vehicle driver does not adjust his or her speed, and that his or her distance has just barely fallen below the safety distance, at time t. In other words, the potential danger of the traffic situation increases, which is illustrated by the rising curve in Fig. 3a.

Consequently, the system issues the motor vehicle driver an action instruction, e.g., in the form of "You must brake" or "Please brake", this action instruction having a higher intonation degree I_2 along with a correspondingly higher connotation level K2. If the motor vehicle driver also does not react to this, then the potential danger of the traffic situation is increased further, which is illustrated by the additional rise in Fig. 3a. This means that a further failure

of the motor vehicle driver to react could lead to an accident in a very short time. This instruction requiring immediate action can, for example, be output in the form of "Brake hard", using command intonation I_3. In this case, the connotation levels illustrated in Fig. 3d represent a climax, i.e. the increase in the expression, from less important to more important. In addition, it should be noted that the changes illustrated in Figs. 2a to 2d and Figs. 3a to 3d are not according to scale, but are rather to be understood as qualitative information.

ABSTRACT

In a method and device for outputting information and/or messages from at least one device using speech, the information and/or messages required for vocal output are provided in a voice memory, the information and/or messages are read by a processing device according to a demand, and the information and/or messages are output via acoustic output device. The information and/or messages are output with a varying intonation according to their relevance.

[11150/30]

METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR STATUS MESSAGES, USING SPEECH

FIELD OF THE INVENTION

The present invention relates to a method and a device for outputting information and/or status messages of at least one electrical device, using speech.

BACKGROUND INFORMATION

Methods and devices of this type are generally used in so-called interactive voice-communication systems or voice-controlled systems for, e.g. vehicles, computers, robots, machines, equipment, etc.

In general, an interactive voice-communication system (SDS) can essentially be reduced to the following components:

- Voice recognition system, which compares an orally input command ("voice command") to other allowed voice commands, and decides which command, in all probability, was orally input;
- Voice output, which outputs the voice commands and signal tones necessary for prompting the user, and possibly acknowledges the recognition result;
- Dialog and sequencing control, in order to explain to the user which type of input is expected, to check if the input is consistent with the prompt and the current status of the application, and to trigger the resulting action in the application (for example, the device to be controlled);
- Control interface as an interface to the application: Hidden behind it are hardware and software modules for controlling various actuators and computers, which contain the application; and
- Application that is controlled by speech: For example, it can be an ordering or information system, a CAE workstation, or a wheel chair for the disabled.

MARKED-UP VERSION

For example, such a voice recognition system is [known from] described in German Published Patent Application No. 195 33 541 [C1]. To increase the acceptance of such man-machine dialog, [the document proposes, for instance, using] synonymous words or various pronunciations for the commands are used, or [else rearranging] the words are rearranged in the commands. For example, "larger radius when turning left" can alternatively be expressed here as "when turning left, larger radius". In addition, [it is suggested that] a multilingual, interactive communication system independent of the speaker can be set up by expanding the memory, it being possible to alternatively switch between the interactive communication systems of various languages. In addition, [the document suggests integrating so-called ellipses, i.e.]ellipses may be used, i.e., dispensing with the repetition of complete command sentences, and instead using commands such as "higher", "sharper", or "further", the voice recognition system then assigning these to the preceding commands. In response to uncertain recognition, the voice recognition system can also pose questions such as "Excuse me?", "Please repeat that", or "What else?", or issue specific suggestions such as "Louder, please". All of these measures are used to avoid monotonic communication[,] and to have the dialog more closely approximate human-to-human communication. To improve the communication, the voice system is coupled to an optical display medium, on which the recognized commands are indicated for control purposes. Furthermore, the optical display medium allows the display of functions from the target device[,] which are set in response to the voice command; and/or the display of various functions/alternatives, which can subsequently be set or selected by a voice command. A disadvantage of [the known] this device and the method implemented [by it] thereby is that, despite the given improvements, the voice output tires the user due to its monotony, so that his or her reaction time is too slow during events requiring immediate action. An additional problem is

that, in response to recognition difficulties, the [known] voice recognition systems [run through] perform an endless loop and issue the user the same prompt again and again, so that the workflow is interrupted.

Therefore, it is an object of the present invention is based on the engineering [problem of providing] to provide a method and a device for outputting information and/or status messages, using speech, in which the attentiveness of the user is improved.

SUMMARY

The above and other beneficial objects [The solution of the engineering problem follows from the features of Claims 1, 9, and 10. Additional advantageous refinements] of the present invention [follow from the dependent claims.] are achieved by providing a device and method as described herein

By using different intonations, the attention of the user is immediately obtained while the speech is being output, so that the reaction time for [carrying out] performing the requested instruction is considerably reduced. In the case of instructions requiring immediate action, the status messages have a command intonation.

To further increase the attention span, and the differentiation of instructions requiring immediate action, the volume of the voice output [can] may be increased [in] for instructions requiring immediate action, and/or these instructions [can] may be inserted in a particularly harsh or abrupt manner.

In addition, the voice recognition system [can] may be designed to be multilingual, so that, for example, one [can] may choose between a man's voice and a woman's voice[; one]. One of these [being] voices are selected by the system, for instructions requiring immediate action, and the other [being]

is selected by the system, for information or status messages not requiring immediate action.

To ensure the workflow, the voice recognition system is only activated by actuating a "Push to talk" [PTT](PTT) switch, the dialog-communication level being changed in the absence of a valid interaction. To increase the recognition reliability and improve the user prompting, individual commands [can] may be saved in various, alternative output forms, which are then successively output in response to an invalid interaction[; the]. The dialog-communication level is only [being] changed when a valid interaction does not ensue in response to all of the command forms. To avoid monotony, the sequence of the output [can] may be permutated by a random-number generator.

The [root idea] basis of the present invention is to use the manner in which speech is output to the motor vehicle driver, in order to create an emotion that causes one to act in accordance with the situation.

[The present invention is explained below in detail, using a preferred exemplary embodiment. The figures show:] BRIEF DESCRIPTION OF THE DRAWINGS

[Fig. 1 a schematic flowchart of the] Figure 1 is a flowchart illustrating a method for automatically controlling at least [one] on device[,] using voice recognition according to the present invention.[;]

[Fig. 2a a schematic characteristic curve of the] Figure 2a illustrates a graph of a potential danger during [interactions not requiring immediate action;
Fig. 2b a corresponding denotation characteristic;
Fig. 2c a corresponding intonation characteristic;
Fig. 2d a corresponding connotation characteristic;
Fig. 3a a schematic characteristic curve of the] an interaction that does not require immediate action.

Figure 2b illustrates a denotation graph corresponding to Figure 2a

Figure 2c illustrates an intonation graph corresponding to Figure 2a

Figure 2d illustrates a connotation graph corresponding to Figure 2a

Figure 3a illustrates a graph of a potential danger during [interactions requiring immediate action;] an interaction that requires immediate action.

[Fig. 3b a corresponding denotation characteristic;] Figure 3b illustrates a denotation graph corresponding to Figure 3a.

[Fig. 3c a corresponding intonation characteristic; and] Figure 3c illustrates an intonation graph corresponding to Figure 3a.

[Fig. 3d a corresponding connotation characteristic.] Figure 3d illustrates a connotation graph corresponding to Figure 3a.

DETAILED DESCRIPTION

The voice recognition system is activated by actuating a PTT switch. For [a better understanding] clarity, the voice output of the voice recognition system is subdivided into commands KOM and prompts Auff which, in reality, [can] may be identical. [From this point on] Hereinafter, commands KOM are to be understood as a direct instruction to act, such as "BRAKE" or "TURN ON LIGHT", whereas prompts Auff request an interaction in the form of an input, such as "Please specify desired temperature in degrees C."

If the voice recognition system now generates a command KOM, then this command KOM is subdivided according to whether it is an instruction requiring immediate action[,] or an instruction

not requiring immediate action. [In simplified terms] More simply, instructions requiring immediate action are commands KOM, which call for the action to be [carried out] performed quickly. An example of this is command KOM "Brake", when an ADR system or a precrash sensory system has detected a collision object. Examples of instructions not requiring immediate action include commands KOM of a navigation system. In this context, instructions requiring immediate action are inserted in time t_1, with command-intonation voice S1 and high volume L1, in a harsh and abrupt manner, in order to produce a high degree of attentiveness in the user. However, instructions not requiring immediate action are inserted softly, at low volume L2 and normal intonation S2.

As a rule, time is not a critical factor in the case of prompts Auff, so that, in this case, good user prompting is [in the forefront] of concern. For this purpose, n different alternatives of a prompt Auff [can] may be stored in the speech memory. For example, the alternatives [can] may be different emphases, pronunciations, word rearrangements, or synonymous terms. After acoustically outputting the first alternative, the voice recognition system waits for a predetermined period of time for an interaction. If no interaction or an invalid interaction [takes place] occurs within this time period, then the voice recognition system repeats the prompt, using the [second] subsequent alternative up to the nth alternative, if necessary. If a valid interaction [takes place] occurs, then this request is [carried out] performed and, if necessary, a new prompt Auff is output. But if no valid interaction occurs [takes place] in response to the nth alternative, then the system switches to another dialog-communication level DKE, in order to ensure the workflow. For example, new dialog-communication level DKE [can] may then be a selection list, which is displayed on the trip-computer monitor, and from which the user [can] may select a corresponding menu.

Figures 2a-d schematically represent the conditions for an instruction not requiring immediate action, such as an information prompt for a navigation system. In Fig. 2a, the importance of the interaction is plotted over time[; instructions]. Instructions for action [being] are output at times t_0 - t_2, and it [being] is assumed that there was no reaction to each preceding prompt. Since a missing input in the navigation system only results in the inoperability of comfort components, which are also not necessarily desired by the motor vehicle driver, the importance does not change over time. The information [about] regarding the content of the command, or the so-called denotation, i.e. the input request, also remains constant over time; [this being represented] as illustrated in Fig. 2b. At time t_0, the motor vehicle driver may be prompted, "Please input your desired destination now." This prompt is issued, using a certain intonation I_1 and a certain connotation K1, which are [represented] illustrated in Figs. 2c [-d] and 2d. If nothing is input, then the system does not know [why this has been omitted, e.g.]the reason for omission, e.g., if the motor vehicle driver did not hear the request or deliberately intended not to [carry] perform it [out]. Therefore, the prompt, "Would you like to input a destination," is issued again at time t_1, using a stronger intonation I_2, in order to improve the possibility of it being perceived. However, connotation level K2 decreases. If, in response, nothing is input again, then the system [can] may certainly [assume] determine that the motor vehicle driver does not wish to do this. To avoid annoying the motor vehicle driver [too much] with constant repetition, a prompt such as "If you do not wish to input a destination, I will now turn myself off" is then issued one last time, at time t_2. This last prompt is output, using a very low intonation I_3, and it just has a low connotation. As [can be gathered from the characteristic] illustrated in Fig. 2d, the connotation forms an anticlimax, i.e., a transition from a strong to a weak expression, whereas a certain variation occurs in the intonation, in order to counteract monotony.

In contrast [to this], Figs. 3a [-d] to 3d illustrate represent a situation in which the importance of the interaction increases over time, until action is finally required. For example, the motor vehicle travels on a motorway[,] at a speed greater than an allowed speed, while maintaining the safety distance behind a motor vehicle. At time t_0, the system issues an action instruction to the motor vehicle driver, e.g., in the form of "Please adjust your speed." The action instruction has a low intonation degree I_1 and a correspondingly low connotation level K1[,] since the motor vehicle driver is indeed acting illegally, but no immediate danger exists. In addition, it is now assumed that the motor vehicle driver does not adjust his or her speed, and that his or her distance has just barely fallen below the safety distance, at time t. In other words, the potential danger of the traffic situation increases, which is [symbolized] illustrated by the rising curve in Fig. 3a.

Consequently, the system issues the motor vehicle driver an action instruction, e.g., in the form of "You must brake" or "Please brake", this action instruction having a higher intonation degree I_2 along with a correspondingly higher connotation level K2. If the motor vehicle driver also does not react to this, then the potential danger of the traffic situation is increased further, which is [represented] illustrated by the additional rise in Fig. 3a. This means that a further failure of the motor vehicle driver to react could lead to an accident in a very short time. This instruction requiring immediate action can, for example, be output in the form of "Brake hard", using command intonation I_3. In this case, the connotation levels [shown] illustrated in Fig. 3d represent a climax, i.e. the increase in the expression, from less important to more important. In addition, it should be noted that the changes [represented] illustrated in Figs. 2a [-d] to 2d and Figs. 3a [-d] to 3d are not according to scale, but are rather to be understood as qualitative information.

ABSTRACT

In a method and device for outputting information and/or messages from at least one device using speech, the information and/or messages required for vocal output are provided in a voice memory, the information and/or messages are read by a processing device according to a demand, and the information and/or messages are output via acoustic output device. The information and/or messages are output with a varying intonation according to their relevance.

[11150/30]

METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR STATUS MESSAGES, USING SPEECH

The present invention relates to a method and a device for outputting information and/or status messages of at least one electrical device, using speech.

Methods and devices of this type are generally used in so-called interactive voice-communication systems or voice-controlled systems for, e.g. vehicles, computers, robots, machines, equipment, etc.

In general, an interactive voice-communication system (SDS) can essentially be reduced to the following components:

- Voice recognition system, which compares an orally input command ("voice command") to other allowed voice commands, and decides which command, in all probability, was orally input;
- Voice output, which outputs the voice commands and signal tones necessary for prompting the user, and possibly acknowledges the recognition result;
- Dialog and sequencing control, in order to explain to the user which type of input is expected, to check if the input is consistent with the prompt and the current status of the application, and to trigger the resulting action in the application (for example, the device to be controlled);
- Control interface as an interface to the application: Hidden behind it are hardware and software modules for controlling various actuators and computers, which contain the application; and
- Application that is controlled by speech: For example, it can be an ordering or information system, a CAE workstation, or a wheel chair for the disabled.

For example, such a voice recognition system is known from German Patent No. 195 33 541 C1. To increase the acceptance of such man-machine dialog, the document proposes, for instance, using synonymous words or various pronunciations for the commands, or else rearranging the words in the commands. For example, "larger radius when turning left" can alternatively be expressed here as "when turning left, larger radius". In addition, it is suggested that a multilingual, interactive communication system independent of the speaker be set up by expanding the memory, it being possible to alternatively switch between the interactive communication systems of various languages. In addition, the document suggests integrating so-called ellipses, i.e. dispensing with the repetition of complete command sentences, and instead using commands such as "higher", "sharper", or "further", the voice recognition system then assigning these to the preceding commands. In response to uncertain recognition, the voice recognition system can also pose questions such as "Excuse me?", "Please repeat that", or "What else?", or issue specific suggestions such as "Louder, please". All of these measures are used to avoid monotonic communication, and to have the dialog more closely approximate human-to-human communication. To improve the communication, the voice system is coupled to an optical display medium, on which the recognized commands are indicated for control purposes. Furthermore, the optical display medium allows the display of functions from the target device, which are set in response to the voice command; and/or the display of various functions/alternatives, which can subsequently be set or selected by a voice command. A disadvantage of the known device and the method implemented by it is that, despite the given improvements, the voice output tires the user due to its monotony, so that his or her reaction time is too slow during events requiring immediate action. An additional problem is that, in response to recognition difficulties, the known voice recognition systems run through an endless loop and issue the user the same prompt again and again, so that the workflow is interrupted.

Therefore, the present invention is based on the engineering problem of providing a method and a device for outputting information and/or status messages, using speech, in which the attentiveness of the user is improved.

The solution of the engineering problem follows from the features of Claims 1, 9, and 10. Additional advantageous refinements of the present invention follow from the dependent claims.

By using different intonations, the attention of the user is immediately obtained while the speech is being output, so that the reaction time for carrying out the requested instruction is considerably reduced. In the case of instructions requiring immediate action, the status messages have a command intonation.

To further increase the attention span, and the differentiation of instructions requiring immediate action, the volume of the voice output can be increased in instructions requiring immediate action, and/or these instructions can be inserted in a particularly harsh or abrupt manner.

In addition, the voice recognition system can be designed to be multilingual, so that, for example, one can choose between a man's voice and a woman's voice; one of these being selected by the system, for instructions requiring immediate action, and the other being selected by the system, for information or status messages not requiring immediate action.

To ensure the workflow, the voice recognition system is only activated by actuating a "Push to talk" PTT switch, the dialog-communication level being changed in the absence of a valid interaction. To increase the recognition reliability and improve the user prompting, individual commands can be saved in various, alternative output forms, which are then

successively output in response to an invalid interaction; the dialog-communication level only being changed when a valid interaction does not ensue in response to all of the command forms. To avoid monotony, the sequence of the output can be permutated by a random-number generator.

The root idea of the present invention is to use the manner in which speech is output to the motor vehicle driver, in order to create an emotion that causes one to act in accordance with the situation.

The present invention is explained below in detail, using a preferred exemplary embodiment. The figures show:

Fig. 1 a schematic flowchart of the method for automatically controlling at least one device, using voice recognition;

Fig. 2a a schematic characteristic curve of the potential danger during interactions not requiring immediate action;

Fig. 2b a corresponding denotation characteristic;

Fig. 2c a corresponding intonation characteristic;

Fig. 2d a corresponding connotation characteristic;

Fig. 3a a schematic characteristic curve of the potential danger during interactions requiring immediate action;

Fig. 3b a corresponding denotation characteristic;

Fig. 3c a corresponding intonation characteristic; and

Fig. 3d a corresponding connotation characteristic.

The voice recognition system is activated by actuating a PTT switch. For a better understanding, the voice output of the voice recognition system is subdivided into commands KOM and prompts Auff which, in reality, can be identical. From this point on, commands KOM are to be understood as a direct instruction to act, such as "BRAKE" or "TURN ON LIGHT", whereas prompts Auff request an interaction in the form of an

input, such as "Please specify desired temperature in degrees C."

If the voice recognition system now generates a command KOM, then this command KOM is subdivided according to whether it is an instruction requiring immediate action, or an instruction not requiring immediate action. In simplified terms, instructions requiring immediate action are commands KOM, which call for the action to be carried out quickly. An example of this is command KOM "Brake", when an ADR system or a precrash sensory system has detected a collision object. Examples of instructions not requiring immediate action include commands KOM of a navigation system. In this context, instructions requiring immediate action are inserted in time t_1, with command-intonation voice S1 and high volume L1, in a harsh and abrupt manner, in order to produce a high degree of attentiveness in the user. However, instructions not requiring immediate action are inserted softly, at low volume L2 and normal intonation S2.

As a rule, time is not a critical factor in the case of prompts Auff, so that, in this case, good user prompting is in the forefront. For this purpose, n different alternatives of a prompt Auff can be stored in the speech memory. For example, the alternatives can be different emphases, pronunciations, word rearrangements, or synonymous terms. After acoustically outputting the first alternative, the voice recognition system waits a predetermined period of time for an interaction. If no interaction or an invalid interaction takes place within this time period, then the voice recognition system repeats the prompt, using the second alternative up to the nth alternative, if necessary. If a valid interaction takes place, then this request is carried out and, if necessary, a new prompt Auff is output. But if no valid interaction takes place

in response to the nth alternative, then the system switches to another dialog-communication level DKE, in order to ensure the workflow. For example, new dialog-communication level DKE can then be a selection list, which is displayed on the trip-computer monitor, and from which the user can select a corresponding menu.

Figures 2a-d schematically represent the conditions for an instruction not requiring immediate action, such as an information prompt for a navigation system. In Fig. 2a, the importance of the interaction is plotted over time; instructions for action being output at times t_0 - t_2, and it being assumed that there was no reaction to each preceding prompt. Since a missing input in the navigation system only results in the inoperability of comfort components, which are also not necessarily desired by the motor vehicle driver, the importance does not change over time. The information about the content of the command, or the so-called denotation, i.e. the input request, also remains constant over time; this being represented in Fig. 2b. At time t_0, the motor vehicle driver may be prompted, "Please input your desired destination now." This prompt is issued, using a certain intonation I_1 and a certain connotation K1, which are represented in Figs. 2c-d. If nothing is input, then the system does not know why this has been omitted, e.g. if the motor vehicle driver did not hear the request or deliberately intended not to carry it out. Therefore, the prompt, "Would you like to input a destination," is issued again at time t_1, using a stronger intonation I_2, in order to improve the possibility of it being perceived. However, connotation level K2 decreases. If, in response, nothing is input again, then the system can certainly assume that the motor vehicle driver does not wish to do this. To avoid annoying the motor vehicle driver too much with constant repetition, a prompt such as "If you do not wish to input a destination, I will now turn myself off" is then issued one last time, at time t_2. This last prompt is output, using a very low intonation I_3, and it just has a low

connotation. As can be gathered from the characteristic in Fig. 2d, the connotation forms an anticlimax, i.e. a transition from a strong to a weak expression, whereas a certain variation occurs in the intonation, in order to counteract monotony.

In contrast to this, Figs. 3a-d represent a situation in which the importance of the interaction increases over time, until action is finally required. For example, the motor vehicle travels on a motorway, at a speed greater than an allowed speed, while maintaining the safety distance behind a motor vehicle. At time t_0, the system issues an action instruction to the motor vehicle driver, e.g. in the form of "Please adjust your speed." The action instruction has a low intonation degree I_1 and a correspondingly low connotation level K1, since the motor vehicle driver is indeed acting illegally, but no immediate danger exists. In addition, it is now assumed that the motor vehicle driver does not adjust his or her speed, and that his or her distance has just barely fallen below the safety distance, at time t. In other words, the potential danger of the traffic situation increases, which is symbolized by the rising curve in Fig. 3a.

Consequently, the system issues the motor vehicle driver an action instruction, e.g. in the form of "You must brake" or "Please brake", this action instruction having a higher intonation degree I_2 along with a correspondingly higher connotation level K2. If the motor vehicle driver also does not react to this, then the potential danger of the traffic situation is increased further, which is represented by the additional rise in Fig. 3a. This means that a further failure of the motor vehicle driver to react could lead to an accident in a very short time. This instruction requiring immediate action can, for example, be output in the form of "Brake hard", using command intonation I_3. In this case, the connotation levels shown in Fig. 3d represent a climax, i.e. the increase in the expression, from less important to more

important. In addition, it should be noted that the changes represented in Figs. 2a-d and Figs. 3a-d are not according to scale, but are rather to be understood as qualitative information.

New Claims 1 through 10

1. A method for outputting information and/or status messages of at least one electrical device, using speech, the information and/or status messages (KOM; Auff) needed for the voice output being stored in a speech memory, read out by a processing device on demand, and output by an acoustic output device, especially a loudspeaker, characterized in that the information and/or status messages (KOM; Auff) are output using different intonations (I), depending on relevance.

2. The method as recited in Claim 1, characterized in that information and/or status messages (KOM) requiring immediate action are output, using command intonation (I).

3. The method as recited in Claim 1 or 2, characterized in that at least information and/or status messages (KOM) requiring immediate action are output at a high volume (L1).

4. The method as recited in one of Claims 1 through 3, characterized in that information and/or status messages (KOM) requiring immediate action are inserted harshly.

5. The method as recited in one of Claims 1 through 4, characterized in that the information and/or status messages (KOM, Auff) are stored in different speaking voices, in the speech memory, and the speaking voice is at least changed in the case of information and/or status messages (KOM) requiring immediate action.

6. The method as recited in one of Claims 1 through 5, characterized in that, at least in the case of information and/or status messages (KOM) requiring

EL594612921US

immediate action, the intonation (I) and the connotation (K) are increased in accordance with importance.

7. The method as recited in one of Claims 1 through 6, characterized in that, in the case of information and/or status messages (KOM) not requiring immediate action, the intonation (I) is varied with decreasing connotation (K).

8. The method as recited in one of Claims 1 through 7, characterized in that it is part of a method for controlling at least one electrical device, using voice recognition.

9. The method as recited in one of the preceding claims, characterized in that various alternatives of the prompts (Auff) are stored in the speech memory; in response to the failure to interact with a prompt (Auff), the alternatives being successively output until an interaction occurs; or in response to the failure to interact with the last alternative, the dialog-communication level (DKE) being changed.

10. A device for outputting information and/or status messages of at least one electrical device via speech, having a speech memory in which the data of the information and/or status messages (KOM) needed for the voice output are stored, and having a processing device and an acoustic output device, characterized in that the information and/or status messages (KOM) can be output using various intonations (I), depending on relevance.

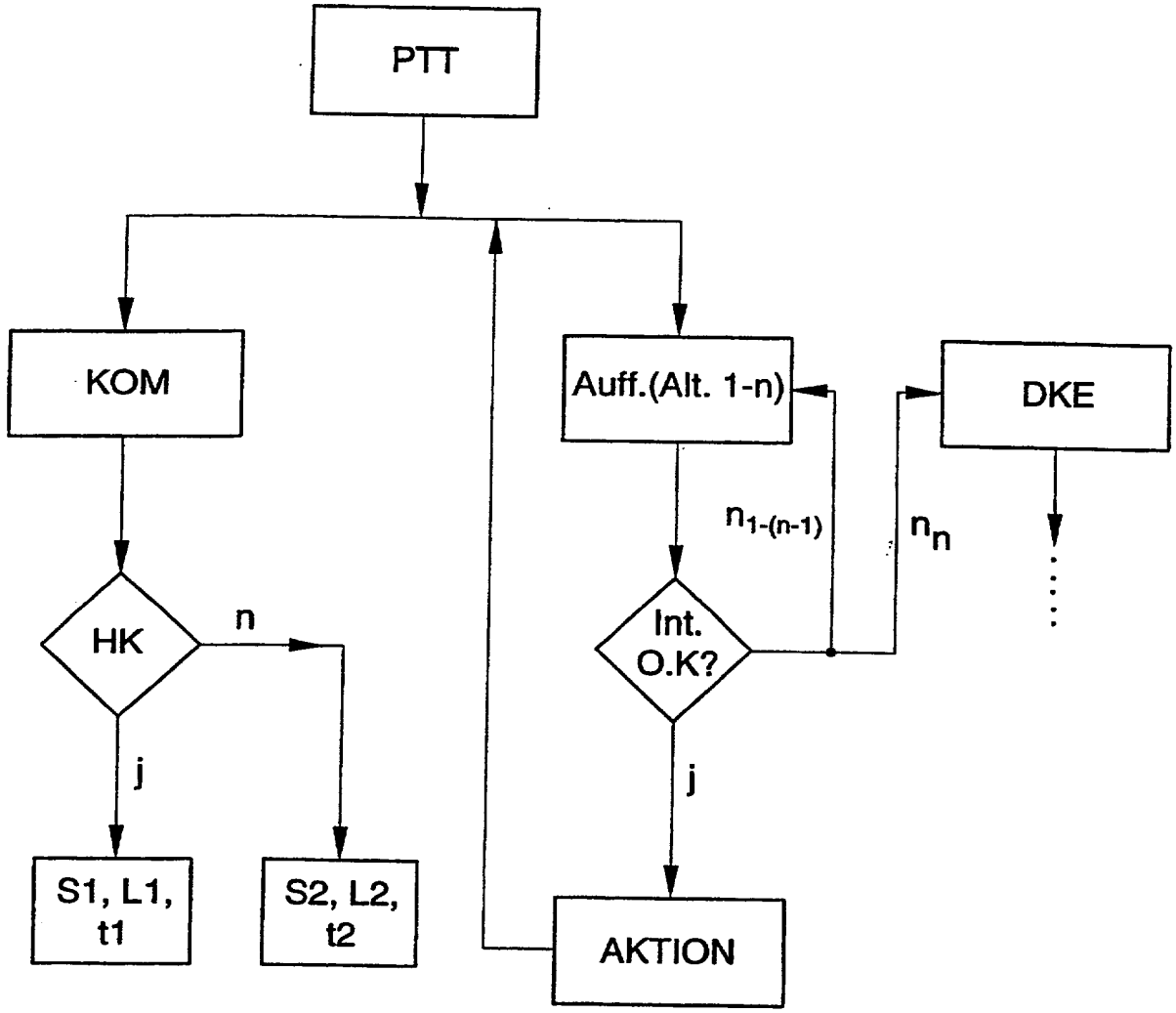
PTT
KOM
Auff.(Alt. 1-n)
DKE
HK
n
j
S1, L1, t1
S2, L2, t2
$n_{1-(n-1)}$
n_n
Int. O.K?
j
AKTION

FIG. 1

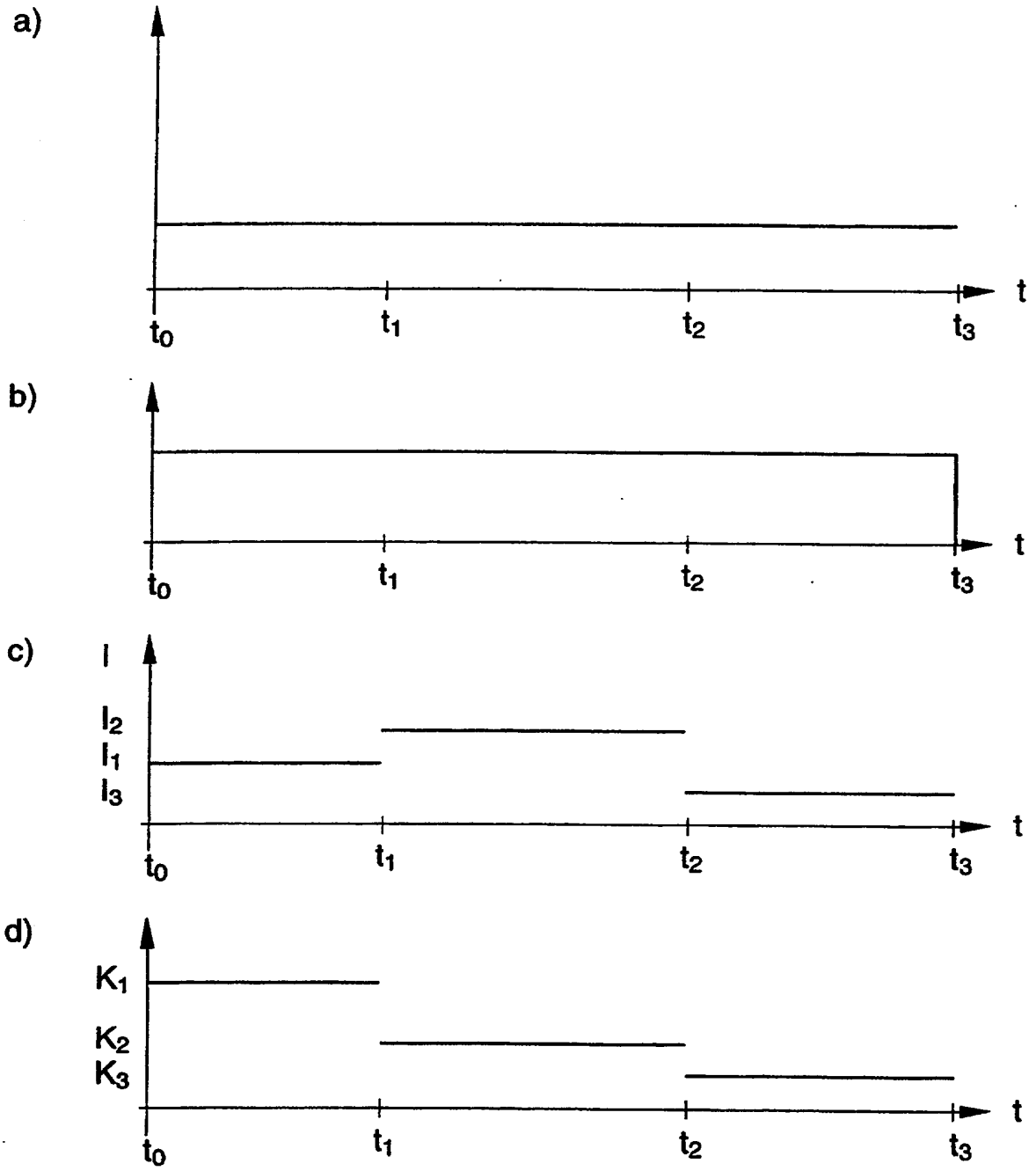
a)
t
t_0
t_1
t_2
t_3
b)
t
t_0
t_1
t_2
t_3
c)
I
I_2
I_1
I_3
t
t_0
t_1
t_2
t_3
d)
K_1
K_2
K_3
t
t_0
t_1
t_2
t_3

FIG. 2

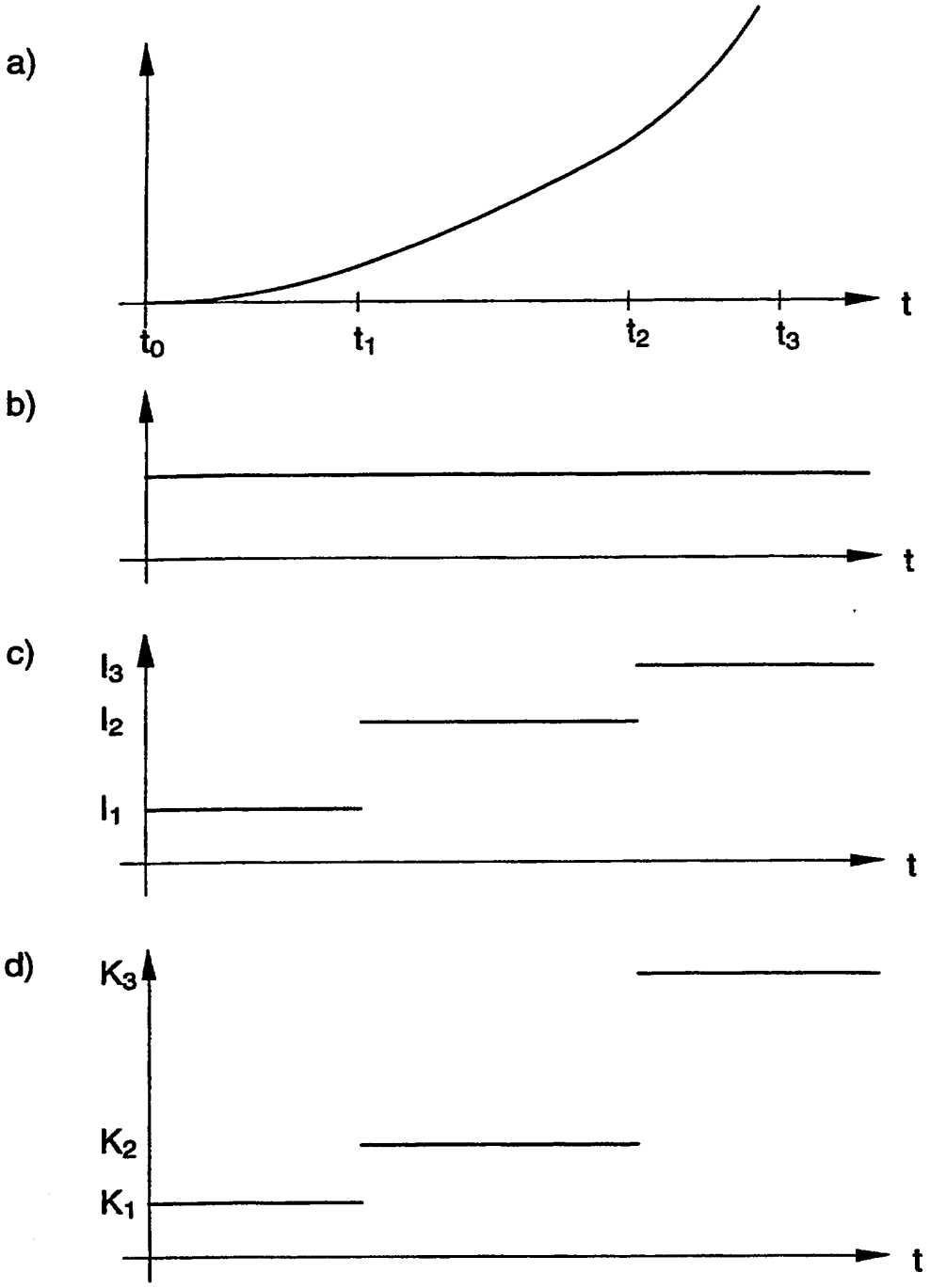

FIG.3

11150/30

COMBINED DECLARATION AND
POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below adjacent to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled **METHOD AND DEVICE FOR OUTPUTTING INFORMATION AND/OR STATUS MESSAGES, USING SPEECH**, and the specification of which:

[ ] is attached hereto;

[ ] was filed as United States Application Serial No. _____ on _____, 19__ and was amended by the Preliminary Amendment filed on _____, 19__.

[X] was filed as PCT International Application Number PCT/EP99/06476 on the 3rd day of September 1999.

[X] an English translation of which is filed herewith.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a). I hereby claim foreign priority benefits under Title 35, United States Code § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international applications(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

EL244504988US

PRIOR FOREIGN/PCT APPLICATION(S)
AND ANY PRIORITY CLAIMS UNDER 35 U.S.C. § 119

Country: Federal Republic of Germany

Application No. 198 47 879.8

Date of Filing: October 16, 1998

Priority Claimed
Under 35 U.S.C. § 119: [X] Yes [ ] No

Country: Federal Republic of Germany

Application No. 199 08 137.9

Date of Filing: February 25, 1999

Priority Claimed
Under 35 U.S.C. § 119: [X] Yes [ ] No

I hereby claim the benefit under Title 35, United States Code § 120 of any United States Application or PCT International Application designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations § 1.56(a) which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

PRIOR U.S. APPLICATIONS OR
PCT INTERNATIONAL APPLICATIONS
DESIGNATING THE U.S. FOR BENEFIT UNDER 35 U.S.C. § 120

U.S. APPLICATIONS

Number :

Filing Date :

PCT APPLICATIONS
DESIGNATING THE U.S.

PCT Number :

PCT Filing Date :

I hereby appoint the following attorney(s) and/or agents to prosecute the above-identified application and transact all business in the Patent and Trademark Office connected therewith.

(List name(s) and registration number(s)):

3

Richard L. Mayer, Reg. No. 22,490
Patrick J. Birde, Reg. No. 29,770
Jeffrey M. Butler, Reg. No. 41,652
_____, Reg. No. _____


26646
PATENT TRADEMARK OFFICE

All correspondence should be sent to:

Richard L. Mayer, Esq.
Kenyon & Kenyon
One Broadway
New York, New York 10004

Telephone No.: (212) 425-7200
Facsimile No.: (212) 425-5288

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-00

Full name of inventor Ingo BOECKMANN

Inventor's signature Böckmann Date 26.03.2001

Citizenship Federal Republic of Germany

Residence Wittenkampeweg 9
D-38551 Ribbesbuettel DEX
Federal Republic of Germany

Post Office Address Same as above

2.00

Full name of inventor Holger EBERT

Inventor's signature Holger Ebert Date 10. April 2001

Citizenship Federal Republic of Germany

Residence Hochstrasse 22
D-90429 Nuernberg DEX
Federal Republic of Germany

Post Office Address Same as above

3-00

Full name of inventor Matthias HEIMERMANN

Inventor's signature M. Heimermann Date 06.04.2001

Citizenship Federal Republic of Germany

Residence Campestrasse 3 A
D-38302 Wolfenbuettel DEX
Federal Republic of Germany

Post Office Address Same as above